

BEZOUT-TYPE THEOREMS FOR DIFFERENTIAL FIELDS

GAL BINYAMINI

ABSTRACT. We prove analogs of the Bezout and the Bernstein-Kushnirenko-Khovanskii theorems for systems of algebraic differential conditions over differentially closed fields. Namely, given a system of algebraic conditions on the first l derivatives of an n -tuple of functions, which admits finitely many solutions, we show that the number of solutions is bounded by an appropriate constant (depending singly-exponentially on n and l) times the volume of the Newton polytope of the set of conditions. This improves a doubly-exponential estimate due to Hrushovski and Pillay.

We illustrate the application of our estimates in two diophantine contexts: to counting transcendental lattice points on algebraic subvarieties of semi-abelian varieties, following Hrushovski and Pillay; and to counting the number of intersections between isogeny classes of elliptic curves and algebraic varieties, following Freitag and Scanlon. In both cases we obtain bounds which are singly-exponential (improving the known doubly-exponential bounds) and which exhibit the natural asymptotic growth with respect to the degrees of the equations involved.

1. INTRODUCTION

In its most elementary form, Bezout's theorem states that a subset of \mathbb{C}^n defined by equations P_1, \dots, P_n of respective degrees d_1, \dots, d_n can have at most $d_1 \cdots d_n$ isolated points. Various generalizations of this statement have been proposed. For example, the Bernstein-Kushnirenko-Khovanskii theorem estimates the number of isolated points in terms of the *mixed volume* of the Newton polytopes $\Delta(P_1), \dots, \Delta(P_n)$. As a consequence of the Bezout theorem and its generalizations, whenever a set defined within the algebraic category happens to be finite, one can produce effective estimates for the size of the set (which often turn out to be fairly accurate).

In this paper we consider generalizations of Bezout's bound for systems of differential equations. The fundamental question is as follows: *given a system of algebraic conditions on an n -tuple of functions and their first l derivatives, which admits finitely many solutions, can one estimate the number of solutions in terms of the degrees of the equations involved?*

This question has been considered by Hrushovski and Pillay in [8]. Their result, quoted in Theorem 2 below, is a powerful analog of the Bezout theorem which similarly allows one to translate qualitative finiteness results obtained using differential-algebraic and model theoretic methods into effective estimates (two examples of a diophantine nature are discussed below).

The explicit estimate in Theorem 2 is stated in terms of slightly different algebraic data than our naive formulation of the Bezout bound, making it difficult

The author was supported by the Banting Postdoctoral Fellowship and the Rothschild Fellowship.

to make a direct comparison. None the less, it is clear that the Bezout bound has two advantages. First, the Bezout bound (and its various generalizations) are singly-exponential with respect to the ambient dimension, whereas Theorem 2 admits doubly-exponential growth. Second, assuming for example that all equations involved have degree d , the Bezout bound is a polynomial in d with exponent equal to the ambient dimension. This is the natural asymptotic for solutions of n -dimensional systems of equations, and is known of course to be optimal in the algebraic context. The bound of Theorem 2 depends polynomially on the degree as well, but the exponent can be in general much larger than the ambient dimension (indeed, exponentially large).

Our principal result is an estimate which recovers the two asymptotic properties above in the differential algebraic context. The statements are presented in §1.2, with formulations in terms of degrees and, more generally, mixed volumes of Newton polytopes.

Theorem 2 has been applied to produce effective bounds for some diophantine counting problems:

- i. In [8] Theorem 2 was used to derive bounds on the number of transcendental lattice points on algebraic subvarieties of semi-abelian varieties. In particular, in the case of the two-dimensional torus this allowed the authors to produce doubly-exponential bounds for a counting problem due to Bombieri, improving the previous repeated-exponential bounds obtained by Buium [5].
- ii. More recently, Theorem 2 has been used in [6] to derive bounds on the number of intersections between isogeny classes of elliptic curves and algebraic varieties. In particular, in the two-dimensional case this allowed the authors to produce doubly-exponential bounds for a counting problem due to Mazur.

Both of these applications follow the same principal strategy, inspired by the work of Buium [2, 3] (see [11, 12] for surveys). Namely, given a counting problem of a diophantine nature (over \mathbb{C} , for example) one first expands the structure by adjoining to \mathbb{C} a differential operator making it into a differentially-closed field with field of constants $\bar{\mathbb{Q}}$. One then writes a system of differential algebraic conditions satisfied by the solutions of the diophantine problem. Finally, one shows (and this is naturally the deeper step involving arguments specific to the problem at hand) that the system, while possibly admitting more solutions than the original diophantine problem, still admits finitely many solutions.

After carrying out the steps above, one can use Theorem 2 to produce an explicit estimate for the number of solutions of the diophantine problem. In §5 we apply our estimates to obtain refined (singly-exponential) bounds for problem (i) and (ii) above.

1.1. Setup and notations. Let K denote a differentially closed field of characteristic zero with differential D , and K_0 its field of constants. Let M denote an ambient space, which we will take to be either an affine space K^n or a Zariski open dense subset thereof. Denote by ξ a coordinate on M . We define the l -th prolongation as $M^{(l)} = M \times K^{nl}$, and denote by $\xi = \xi^{(0)}, \xi^{(1)}, \dots, \xi^{(l)}$ coordinates on $M^{(l)}$. For simplicity of the notation we denote $s := \dim M^{(l)} = n(l+1)$. We denote by

$\pi_M : M^{(l)} \rightarrow M$ the projection onto M . If there is no ambiguity we omit M from this notation.

For any $x \in M$, denote by $x^{(l)} \in M^{(l)}$ the l -jet of x , i.e. the element $(x, \dots, D^l x)$. We denote by $\mathcal{J}^l(M) \subset M^{(l)}$ the set of all l -jets in $M^{(l)}$,

$$\mathcal{J}^l(M) := \{x^{(l)} : x \in M\}. \quad (1)$$

Let $\Omega^{(l)} \subset M^{(l)}$ denote a Zariski open dense subset of $M^{(l)}$ (we include l in the notation $\Omega^{(l)}$ in the interest of clarity, but this is in fact an arbitrary open dense set). If we do not explicitly specify $\Omega^{(l)}$ then it is taken to be equal to $M^{(l)}$.

Let $Y \subset \Omega^{(l)}$ be a Zariski closed set. We view Y as a system of algebraic-differential conditions on M . We say that $x \in M$ is a *solution* for Y if $x^{(l)} \in Y$. We define the *reduction* of Y $\mathcal{R}(Y) \subset Y$ to be

$$\mathcal{R}(Y) := \text{Clo}[\mathcal{J}^l(M) \cap Y] \quad (2)$$

where Clo denotes Zariski closure. In other words, the reduction of Y is the Zariski closure of the set of jets of the solutions of Y . If Y admits finitely many solutions then $\mathcal{R}(Y)$ is the set consisting of their jets. In this case we denote the number of solutions by $\mathcal{N}(Y)$.

Example 1. Suppose $M = K^2$ with coordinates ξ, η , and consider the set $Y \subset M^{(1)}$ given by the conditions $\xi\eta = 1$ and $\xi^{(1)}\eta + \xi\eta^{(1)} = 1$. Then $\dim Y = 2$. However, it is clear that Y admits no solutions, since deriving the first condition we see that any solution must also satisfy $\xi^{(1)}\eta + \xi\eta^{(1)} = 0$, contradicting the second condition. Thus $\mathcal{R}(Y) = \emptyset$.

In [8], Hrushovski and Pillay considered the problem of estimating $\mathcal{N}(Y)$ in terms of the algebraic degree of Y . In particular, using differential algebraic methods they have established the following estimate.

Theorem 2 ([8, Proposition 3.1]). Let $X \subset M$ and $S \subset \Omega^{(l)}$ be Zariski closed. Denote $m := \dim X$. Denote $Y = S \cap \pi^{-1}(X)$ and suppose that Y admits finitely many solutions. Then

$$\mathcal{N}(Y) \leq \deg(X)^{l2^{ml}} \deg(S)^{2^{ml}-1}. \quad (3)$$

1.2. Statement of our results. Recall that for a polynomial $P \in K[\xi^{(0)}, \dots, \xi^{(l)}]$, its *Newton polytope* denoted $\Delta(P)$ is defined to be the convex hull of the set of exponents (of terms with non-zero coefficients) of P , viewed as a subset of \mathbb{R}^s . We say that Δ is a *lattice polytope* if it is obtained as the Newton polytope of some polynomial. We will denote by $\Delta_{\xi^{(j)}}$ the standard simplex in the $\xi^{(j)}$ variables, and by $\Delta_{\xi}^{(j)}$ the standard simplex in the $\xi^{(0)}, \dots, \xi^{(l)}$ variables.

The classical Bernstein-Kushnirenko-Khovanskii theorem (henceforth BKK) relates the number of solutions of a system of polynomial equations to the mixed volume of their Newton polytopes (we refer the reader to [1] for the notion of mixed volume).

Theorem 3 ([10, 1]). Let $\Delta_1, \dots, \Delta_s$ be lattice polytopes in \mathbb{R}^s . Then for any tuple P_1, \dots, P_s with $\Delta(P_i) = \Delta_i$, the system of equations $P_1 = \dots = P_n = 0$ admits at most μ isolated solutions in $(K^*)^s$, where

$$\mu = n!V(\Delta_1, \dots, \Delta_s) \quad (4)$$

and $V(\dots)$ denotes the mixed volume. Moreover, for a sufficiently generic choice of the tuple P_i the bound μ is attained.

Remark 4. Under some mild conditions, the BKK estimate also bounds the number of zeros in K^s . For instance, this is true if Δ is the convex hull of any finite co-ideal $I \subset \mathbb{Z}_{\geq 0}$ (since in this case one can always, after a generic translation, assume that none of the solutions lie on the coordinate axes). We will refer to such polytopes as co-ideal polytopes. The common case of the polytope of polynomials with bounded degree or multi-degree certainly satisfies this condition.

Our results can be applied either to $(K^*)^s$ or to K^s , but in the latter case we tacitly assume that all polytopes involved are co-ideal polytopes.

Our main result is an analog of the first part of the BKK theorem. Namely, for a system of algebraic-differential conditions Y admitting finitely many solutions, we estimate the number of solutions in terms of a mixed volume of the Newton polytopes associated to the equations defining Y . Note that our various formulations are stated in terms of the degrees of the equations defining Y , whereas Theorem 2 is stated in terms of the algebraic degree of Y as a variety.

We begin with a formulation valid for complete intersections (for a slightly more general form valid for flat limits of complete intersections see Proposition 27).

Theorem 5. Let $Y \subset \Omega^{(l)}$ be a complete intersection defined by polynomials with Newton polytopes $\Delta_1, \dots, \Delta_k$, and suppose that $N(Y) < \infty$. Denote

$$\Gamma := (s+1)\Delta_\xi + \Delta_1 + \dots + \Delta_k. \quad (5)$$

Then

$$N(Y) \leq C_{s,k} V(\Delta_1, \dots, \Delta_k, \Gamma, \dots, \Gamma) \quad C_{s,k} = (s!)(\delta+2)^{\delta(\delta+1)/2} \quad (6)$$

where $\delta := s - k$.

We note that the factor of $s!$ in $C_{s,k}$ is the usual factor appearing in the BKK theorem. The additional factor is an artifact of our construction and could certainly be improved somewhat.

We next present a result valid for arbitrary varieties rather than complete intersections.

Theorem 6. Let $Y \subset \Omega^{(l)}$ be any variety with $N(Y) < \infty$. Suppose that

- (1) Y is contained in a complete-intersection defined by polynomials with Newton polytopes $\Delta_1, \dots, \Delta_k$.
- (2) Y is set-theoretically cut out by equations with Newton polytope Δ containing $\Delta_\xi^{(l)}$. To simplify the notation we write $\Delta_j = \Delta$ for $j > k$.

Denote

$$\Gamma_j := (s+1)\Delta_\xi + \Delta_1 + \dots + \Delta_j. \quad (7)$$

Then

$$N(Y) \leq \sum_{j=k}^s C_{s,j} V(\Delta_1, \dots, \Delta_j, \Gamma_j, \dots, \Gamma_j). \quad (8)$$

In particular, assuming $\Delta_j \subset \Delta$ for $j = 1, \dots, k$, we have

$$N(Y) \leq E_{s,k} V(\Delta_1, \dots, \Delta_k, \Delta, \dots, \Delta), \quad E_{s,k} = \sum_{j=k}^s (2s)^{s-j} C_{s,k}. \quad (9)$$

As an immediate consequence we obtain the following analog of the Kushnirenko theorem.

Theorem 7. *Let $Y \subset \Omega^{(l)}$ be any variety of top dimension $s-k$ cut out by equations with a given Newton polytope Δ containing $\Delta_\xi^{(l)}$. Suppose that Y admits finitely many solutions. Then*

$$\mathcal{N}(Y) \leq E_{s,k} \text{Vol}(\Delta) \quad (10)$$

More generally, for systems admitting infinitely many solutions we have similar estimates for the degree of their reduction. Here the degree of an irreducible variety in $\Omega^{(l)}$ is defined to be the number of intersections with a generic affine-linear space of complementary dimension, and this notion is extended by linearity to arbitrary (possible mixed-dimensional) varieties. We give an analog of the second part of Theorem 6, but the other statements extend in a similar manner.

Corollary 8. *Let $Y \subset \Omega^{(l)}$ and suppose that*

- (1) *Y is contained in a complete-intersection defined by polynomials with Newton polytopes $\Delta_1, \dots, \Delta_k$.*
- (2) *Y is set-theoretically cut out by equations with Newton polytope Δ containing $\Delta_\xi^{(l)}$.*

Assume further that $\Delta_j \subset \Delta$ for $j = 1, \dots, k$. Then

$$\deg(\mathcal{R}(Y)) \leq (s-k+1)E_{s,k}V(\Delta_1, \dots, \Delta_k, \Delta, \dots, \Delta). \quad (11)$$

Finally, we record a corollary of Theorem 7 with a formulation more similar to that of Theorem 2. In particular, it shows that for a fixed system of differential conditions S , the number of solutions within a variety X defined by equations of degree d_X grows asymptotically like d_X^n , which is the expected order of growth (even in a purely algebraic context).

Corollary 9. *Let $Y \subset \Omega^{(l)}$ be a variety of top dimension $s-k$ cut out by equations of degree d_S and $X \subset M$ be a variety cut out by equations $d_X \geq d_S$. Denote $Y = S \cap \pi^{-1}(X)$ and suppose that Y admits finitely many solutions. Then*

$$\mathcal{N}(Y) \leq E_{s,k} d_X^n d_S^{nl} \quad (12)$$

1.3. Elementary differential algebraic constructions. For any Zariski closed set $V \subset M$ we define the l -th prolongation $V^{(l)} \subset M^{(l)}$ to be the Zariski closure of $\{x^{(l)} : x \in V\}$. The prolongation of a finite union of closed sets is clearly equal to the union of their prolongations. Moreover, by a theorem of Kolchin [9], the prolongation of an irreducible set is itself irreducible.

The following encapsulates a key property of differentially closed fields.

Fact 10 ([8, Fact 3.7]). *Let V denote an irreducible variety and $W \subset V^{(1)}$ an irreducible variety which projects dominantly on V . Then for any non-empty Zariski open subset $U \subset W$ there exists $x \in V$ with $x^{(1)} \in U$.*

We denote $M^{(l)(1)} := (M^{(l)})^{(1)}$ and similarly for $x^{(l)(1)}$. We have a canonical variety $\Xi \subset M^{(l)(1)}$ defined to be the Zariski closure of $\{x^{(l)(1)} : x \in M\}$. It is given by the equations $\xi^{(k)(0)} = \xi^{(k-1)(1)}$ for $k = 1, \dots, l$.

Proposition 11. *The points $y \in M^{(l)}$ that satisfy $y^{(1)} \in \Xi$ are precisely the points of the form $x^{(l)}$ for some $x \in M$.*

Proof. Let $y = (x_0, x_1, \dots, x_l)$ with $y^{(1)} = (y, Dy) \in \Xi$. It follows that $x_k = Dx_{k-1}$ for $k = 1, \dots, l$, so $y = (x_0)^{(l)}$ as claimed. The other direction is obvious. \square

Lemma 12. *Let $Y \subset \Omega^{(l)}$ be a variety admitting finitely many solutions. Then $Y^{(1)} \cap \Xi$ does not project dominantly on any positive-dimensional component of Y . More generally, the same holds if $\mathcal{R}(Y)$ does not contain any component of Y .*

Proof. It is enough to check the case of Y irreducible (and positive dimensional). Let $U = \Omega^{(l)} \setminus \mathcal{R}(Y)$. By assumption, $U \cap Y$ is open dense in Y .

Assume toward contradiction that $Y^{(1)} \cap \Xi$ projects dominantly on Y . Then some irreducible component W of this intersection projects dominantly on Y as well. Thus we can apply Fact 10 to the non-empty open subset $W \cap \pi^{-1}(U)$ and deduce that there exists $y \in U$ with $y^{(1)} \in Y^{(1)} \cap \Xi$. But by Proposition 11 such y is of the form $z^{(l)}$ for some $z \in M$, and hence z is another solution of Y , contradicting the definition of U . \square

1.4. Overview of the proof.

1.4.1. *The reduction of dimension.* Let $Y \subset \Omega^{(l)}$ be a variety of positive dimension admitting finitely many solutions, and let $\tilde{Y} := \pi_{\Omega^{(l)}}(Y^{(1)} \cap \Xi)$. Then by Lemma 12 we have $\dim \tilde{Y} < \dim Y$. If $x \in M$ is any solution of Y , then $x^{(l)} \in Y$ and hence $x^{(l)(1)} \in Y^{(1)} \cap \Xi$, so $x^{(l)} \in \tilde{Y}$. That is, x is also a solution of \tilde{Y} . Since $\tilde{Y} \subset Y$ we conclude that $\mathcal{N}(Y) = \mathcal{N}(\tilde{Y})$.

Repeating this reduction s times, one is eventually reduced to counting the number of solutions of a zero-dimensional variety, which is certainly bounded by the number of points in the variety. This is similar to the approach employed in [8]. In order to obtain good effective estimates it is thus necessary to have an effective description of \tilde{Y} in terms of Y , and the key step is obtaining an effective description of $Y^{(1)}$.

1.4.2. *An approximation for the first prolongation.* We now work with an arbitrary ambient space N , and the reader should keep in mind that eventually we will take this ambient space to be $M^{(l)}$ or its open dense subset $\Omega^{(l)}$. To avoid confusion we denote the coordinates on N by ζ .

If $V \subset N$ is an effectively smooth complete intersection defined by k polynomial equations $\{P_j\}$ with a given Newton polytope Δ , then one can explicitly write $2k$ equations in $\zeta, \zeta^{(1)}$ for $V^{(1)}$, with (essentially) the same Newton polytope in ζ and linear in $\zeta^{(1)}$. However, this system of equations degenerates if V is a non-smooth intersection.

Our proof is based on the following idea. We embed V as the zero fiber $V = X_0$ of a flat family X whose generic fiber is a complete intersection. Making a small perturbation we may also assume that the generic fiber is smooth. We write the system of $2k$ equations as above (now depending on an extra deformation parameter e), obtaining a family $\tau(X)$. It turns out that the limit $\tau(X)_0$, while not necessarily equal to $V^{(1)}$, still approximates it rather well: the two agree over an open dense set. To conclude, if V is a limit of a family of complete intersections of a given Newton polygon, then the same is essentially true for $V^{(1)}$ (at least in an open dense set).

1.4.3. Conclusion. Returning now to the notations of §1.4.1, we show that if Y was the limit of a family of complete intersections with a given Newton polytope, then the same is true for \tilde{Y} (with a slightly larger Newton polytope). We may now repeat the construction s times and eventually obtain the limit of a zero-dimensional variety. Of course, the number of points of such a limit is bounded by the number of points of the generic fiber, which may now be estimated with the help of the BKK theorem.

1.5. Acknowledgements. I would like to express my gratitude to Ehud Hrushovski for suggesting the problem studied in this paper to me and for invaluable discussions. I also wish to thank Omar Leon Sanchez for his useful suggestions during the preparation of this manuscript.

2. SMOOTH APPROXIMATIONS AND FLAT FAMILIES

In this section we denote the ambient space by N and its dimension by n . The reader should keep in mind that eventually we will take this ambient space to be $M^{(l)}$ or its open dense subset $\Omega^{(l)}$. To avoid confusion we denote the coordinates on N by ζ .

2.1. The τ -variety associated to a flat family. Let $V \subset N$ be given by

$$V = \{P_1 = \cdots = P_k = 0\}, \quad P_j \in K[\zeta]. \quad (13)$$

We define $\tau(V) \subset N^{(1)}$ to be

$$\tau(V) := \{(\zeta, \zeta^{(1)}) : P_j(\zeta) = 0, (dP_j)_\zeta(\zeta^{(1)}) + P_j^D(\zeta) = 0, j = 1, \dots, k\} \quad (14)$$

where P^D is the polynomial obtained by applying D to each coefficient of P . It follows from the chain rule that $V^{(1)} \subset \tau(V)$. Moreover, if V is effectively smooth then equality holds [8, Remark 3.5.(3)], and in particular $\tau(V)$ has pure codimension $2k$ in $N^{(1)}$.

Recall that a variety $X \subset N \times \mathbb{A}_K$ is flat over \mathbb{A}_K if and only if every component of X projects dominantly on \mathbb{A}_K . For general X , we denote by $\mathcal{F}(X)$ the flat family obtained by removing any component which projects to a point in \mathbb{A}_K . For any $\varepsilon \in \mathbb{A}_K$ we denote by X_ε the ε -fiber of X .

Consider now a variety $X \subset N \times \mathbb{A}_K$ given by

$$X = \mathcal{F}[\{P_1 = \cdots = P_k = 0\}], \quad P_j \in K[\zeta, e] \quad (15)$$

where e denotes a coordinate on \mathbb{A}_K . We say that X is a *generic complete intersection* if the generic fiber X_ε has pure codimension k . We define $\tau(X) \subset N^{(1)} \times \mathbb{A}_K$ to be

$$\tau(X) := \mathcal{F}[\{(\zeta, \zeta^{(1)}) : P_j(\zeta) = 0, (dP_j)_\zeta(\zeta^{(1)}) + P_j^D(\zeta) = 0, j = 1, \dots, k\}]. \quad (16)$$

For generic $\varepsilon \in K_0$, $\tau(X)_\varepsilon$ is just $\tau(X_\varepsilon)$. The following subsection establishes a precise sense in which $\tau(X)_0$, obtained as the limit of these generic fibers, approximates $(X_0)^{(1)}$.

2.2. Approximation of the first prolongation. In this section it will be convenient for us to assume that the field K has an analytic realization. We therefore assume that the field of constants K_0 is a subfield of \mathbb{C} , and that any all functions involved in the definition of any of the varieties we consider have been embedded in the field of meromorphic functions on the disc \mathbb{D} (this is always possible by a result of Seidenberg [13, 14]). Thus we may consider K -varieties as analytic sets.

We begin with a simple lemma.

Lemma 13. *Let $X \subset N \times \mathbb{A}_K$ be a flat family, and $x \in X_0$. Then for any sequence $\varepsilon_i \in K_0$ with $\varepsilon_i \rightarrow 0$, there exists a sequence of K -points $x_i \in X_{\varepsilon_i}$, holomorphic in a common disc and converging uniformly to x .*

Proof. Intersecting with generic linear functionals vanishing at x one can reduce the problem to the case that X is a curve. Moreover, changing coordinate $e \rightarrow e^\nu$ we may assume that the curve is irreducible, smooth and transversal to e at x in the K -sense, i.e. for generic t . We restrict the disc \mathbb{D} to make this true for every t .

Consider the intersection $X \cap \{e = \varepsilon_i\}$. This is a zero dimensional set, and for sufficiently small ε_i contains precisely one solution x_i near x . All x_i are K -points and moreover, are in fact defined over the field of definition of X . Therefore they may be viewed as (a-priori, ramified) holomorphic functions on \mathbb{D} . But in fact $x_i(t)$ is well-defined for $t \in \mathbb{D}$, and thus no ramification can occur and the functions $x_i(t)$ are holomorphic in \mathbb{D} . Finally, x_i converges pointwise to x by definition, and it follows by standard arguments that convergence is uniform (perhaps on a smaller disc). \square

We remark that since the Zariski topology is coarser than the analytic \mathbb{C} -topology, the converse also holds: if $x_i \in X_{\varepsilon_i}$ is a sequence of K points holomorphic on a disc \mathbb{D} and converging uniformly to $x \in K$, then any Zariski closed set containing x_i for every i must contain x .

Proposition 14. *Let X be a generic complete intersection as in (15). Then we have $(X_0)^{(1)} \subset \tau(X)_0$.*

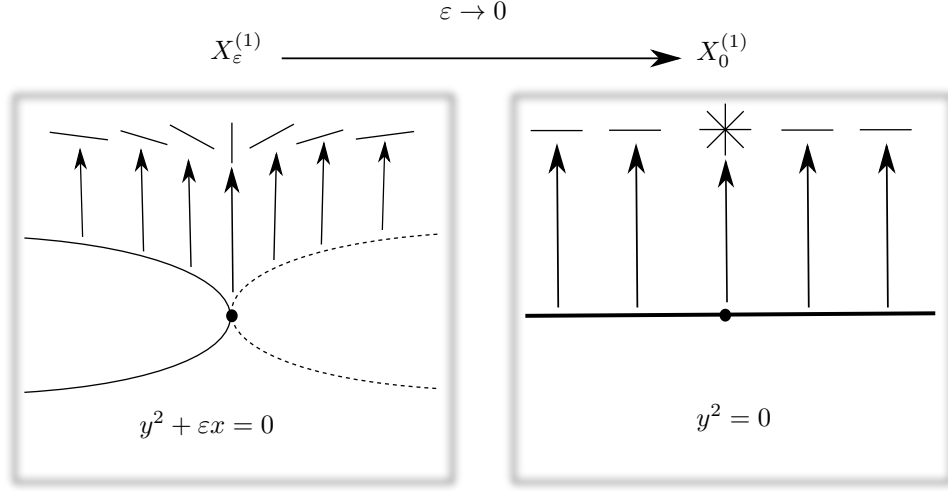
Proof. Since $\tau(X)_0$ is closed by definition, it is enough to prove that for every $x \in X_0$, we have $x^{(1)} \in \tau(X)_0$. Let $x \in X_0$. Since the family X is flat, we can choose by Lemma 13 a sequence $x_i \in X_{\varepsilon_i}$ with $\varepsilon_i \in K_0$ and $\varepsilon_i \rightarrow 0$, converging uniformly to x . Then $(x_i)^{(1)}$ converges uniformly to $x^{(1)}$ (in an appropriate disc). Since $(X_{\varepsilon_i})^{(1)} \subset \tau(X_{\varepsilon_i})$ and $\tau(X_{\varepsilon_i}) = \tau(X)_{\varepsilon_i}$ by the remark following (16), we have $(x_i)^{(1)} \in \tau(X)_{\varepsilon_i}$. Since $\tau(X)$ is closed we conclude that $x^{(1)} \in \tau(X)_0$ as claimed \square

For a variety $V \subset N$ as in (13), we can describe $\tau(V)$ at effectively smooth points in terms of complex tangent spaces. Indeed, let $x \in V$ be an effectively smooth K -point. Then V is smooth as a complex variety at $x(t)$ for generic t and we have the equivalence

$$(x, v) \in \tau(V) \iff \partial_t + v \cdot \partial_\zeta \in T_x V \quad (17)$$

where the right hand side is understood to hold for generic t . In particular, if V is any variety then one can choose a system of equations such that it becomes effectively smooth at every smooth point. Thus, since we know that $\tau(V)$ agrees with $V^{(1)}$ over the effectively smooth locus, we conclude that for every smooth point $x \in V$ we also have

$$(x, v) \in V^{(1)} \iff \partial_t + v \cdot \partial_\zeta \in T_x V. \quad (18)$$

FIGURE 1. Degeneration of $\tau(X)$ in Example 16

The following proposition establishes a partial converse to Proposition 14.

Proposition 15. *Let X be a generic complete intersection as in (15) such that the generic fiber X_ε is effectively smooth. Then there is an open dense set $U \subset X_0$ such that*

$$(X_0)^{(1)} \cap \pi^{-1}(U) = \tau(X)_0 \cap \pi^{-1}(U). \quad (19)$$

Before proving this proposition, we illustrate with the following example.

Example 16. *Consider $X \subset \mathbb{A}_K^2 \times \mathbb{A}_K$ given by $y^2 = ex$. Since the family is defined by an equation with constant coefficients, $(X_0)^{(1)}$ agrees with the tangent bundle of X_0 . For generic ε , the fiber X_ε is an effectively smooth parabola, and for ε constant $\tau(X)_\varepsilon$ agrees with the tangent bundle of this parabola.*

The fiber $\tau(X)_0$ is the flat limit of $\tau(X)_\varepsilon$. Since ∂_y is tangent to X_ε at $(x, y) = (0, 0)$ for any ε , the fiber of $\tau(X)_0$ over this point contains ∂_y . Thus $\tau(X)_0$ is strictly larger than the tangent bundle of X_0 . However, over any point of $U := X_0 \setminus \{(0, 0)\}$, can easily check that the limit of X_ε agrees with the tangent bundle of X_0 .

Proof of Proposition 15. To avoid confusion, in this proof we make the convention that p denotes \mathbb{C} -points, whereas x denotes K -points.

Since the map e is flat on X (over K), it follows that for generic t there are no components of X contained in the $e = 0$ fiber. After restricting to a disc where this holds, the map $e : X \rightarrow \mathbb{C}$ can be viewed as a flat map over \mathbb{C} . By a theorem of Hironaka [7], there exists an analytic stratification $\{Z_\alpha\}$ of X_0 with the following property: *for any sequence of points $p_i \in X$ converging to $p \in Z_\alpha$, if $X_{e(p_i)}$ is smooth at p_i and $T_{p_i}X_{e(p_i)}$ converges to a limit T , then $T_p Z_\alpha \subset T$.* This is somewhat weaker form of Thom's A_e -condition, which will suffice for our purposes. We fix such a stratification, and let Z_0 denote the (union of) top-dimensional stratas.

The following observation is the key geometric idea for the proof. Let $p_i \in X$ be any sequence of points converging to $p \in Z_0$ with $X_{e(p_i)}$ smooth at p_i and $v_i \in T_{p_i}X_{e(p_i)}$ a sequence of tangent vectors converging to a vector v . Then $v \in T_p X_0$. Indeed, one can always pass to a subsequence such that $T_{p_i}X_{e(p_i)}$ converges to some

limit T , and hence $T_p Z_0 \subset T$. But since the dimensions of these sets agree, in fact $T_p X_0 = T_p Z_0 = T$, and in particular $v \in T$ implies $v \in T_p X_0$.

We know by Proposition 14 that $(X_0)^{(1)} \subset \tau(X)_0$. The set of points in X_0 such that the π -fibers of these two sets are equal is K -constructible. Thus, if it does not contain an open dense set it must be contained in a closed set of strictly smaller dimension. Therefore it will suffice to establish the claim over any *analytic* open dense set U . We will establish it with $U = Z_0$. Thus, let $(x, v) \in \tau(X)_0$ with $x \in Z_0$, and we will prove that $(x, v) \in (X_0)^{(1)}$.

Since $\tau(X)$ is flat by definition, we may by Lemma 13 choose a sequence $\varepsilon_i \in K_0$ with $\varepsilon_i \rightarrow 0$ and a sequence $(x_i, v_i) \in \tau(X)_{\varepsilon_i}$ such that (x_i, v_i) are defined in a common disc \mathbb{D} and converge uniformly to (x, v) . By assumption, we may take the fibers X_{ε_i} to be effectively smooth. Also, by the remark following (16) we may assume that $\tau(X)_{\varepsilon_i} = \tau(X_{\varepsilon_i})$. We conclude from (17) that

$$\partial_t + v_i \cdot \partial_\zeta \in T_{x_i} X_{\varepsilon_i} \quad (20)$$

for generic t .

For t outside a countable set, each fiber X_{ε_i} is effectively smooth (over \mathbb{C}) at $x_i(t)$. By the key geometric observation above it now follows that

$$\partial_t + v \cdot \partial_\zeta \in T_x X_0 \quad (21)$$

for generic t . Thus (18) gives $(x, v) \in (X_0)^{(1)}$ as claimed. \square

Corollary 17. *Let X be a generic complete intersection as in (15) and suppose that the generic fiber X_ε is effectively smooth. Then every component of $(X_0)^{(1)}$ is a component of $\tau(X)_0$.*

Proof. Indeed, if Z_1, \dots, Z_r denote the components of X_0 then they each have codimension k , and $Z_i^{(1)}$ are the components of $(X_0)^{(1)}$. In particular, $\pi^{-1}(U)$ is dense in each $Z_i^{(1)}$ (for U given in Proposition 15), and since $\tau(X)_0$ agrees with $(X_0)^{(1)}$ over $\pi^{-1}(U)$ it follows that each $Z_i^{(1)}$ is also a component of $\tau(X)_0$. \square

3. CONSTRUCTIONS WITH FLAT FAMILIES

Once again, in this section we denote the ambient space by N and its dimension by n . The reader should keep in mind that eventually we will take this ambient space to be $M^{(l)}$ or its open dense subset $\Omega^{(l)}$. To avoid confusion we denote the coordinates on N by ζ .

When speaking about the Newton polytope of a polynomial in $K[\zeta, e]$ we mean the Newton polytope in the ζ variables obtained for a generic value of e .

3.1. General lemmas on perturbations and intersections in flat families.

The following proposition shows that the 0-fiber of a flat family is not changed if one makes a sufficiently small perturbation of the family, i.e. a perturbation of sufficiently high order in e .

Proposition 18. *Let $X \subset N \times \mathbb{A}_K^2$ be a variety, and suppose that every component of V projects dominantly on \mathbb{A}_K^2 . Then the flat limit at the origin of $X_{s,t}$ (as a variety) along $\Gamma = \{t = 0\} \subset \mathbb{A}_K^2$ is the same as the flat limit at the origin along any smooth curve $\Gamma' \subset \mathbb{A}_K^2$ sufficiently tangent to Γ .*

Proof. We may assume that X is irreducible. Denote the projection to \mathbb{A}_K^2 by η . Recall that N is an open dense subset of some projective space $\mathbb{C}P^S$. Let \tilde{X} denote the closure of X in $\mathbb{C}P^S \times \mathbb{A}_K^2$, which clearly projects dominantly to \mathbb{A}_K^2 as well.

Consider the map Φ taking a pair (s, t) to the fiber $X_{s,t} := \tilde{X} \cap \eta^{-1}(s, t)$ with its cycle structure in the Chow variety parametrizing cycles of degree $\deg X$ in $\mathbb{C}P^S$. This is a rational map defined over the (open dense) locus $U = \mathbb{A}_K^2 \setminus \Sigma$ such that the intersection above is proper. We have $\text{codim } \Sigma \geq 2$ (otherwise $\dim \eta^{-1}(\Sigma) = \dim X$, which is ruled out by the hypotheses). Thus Φ is defined on both Γ and Γ' except for possibly finitely many exceptions where they intersect Σ .

In projective coordinates on the Chow variety, we have

$$\Phi(s, 0) = s^\nu [\tilde{X}] + O(s^{\nu+1}) \quad (22)$$

where $[\tilde{X}]$ denotes the Chow form of the flat limit of $X_{s,t}$ along Γ (with its cycle structure). Thus

$$\Phi(s, t) = s^\nu [\tilde{X}] + O(s^{\nu+1}) + O(t). \quad (23)$$

We see that if $t = O(s^{\nu+1})$ along Γ' then the limit of $X_{s,t}$ along Γ' is also equal to $[\tilde{X}]$ as a cycle. In particular, these cycles intersect N in the same set, proving the claim. \square

Remark 19. Let U denote a Zariski open subset in the space of n -tuples of polynomials with a given Newton polytope. Proposition 18 implies, in particular, that one may deform a given complete intersection family (15) to make $P_i(\zeta, e) \in U$ for generic values of e , without changing the fiber X_0 . Indeed, consider $\tilde{P}_i = P_i + sQ_i$ where Q_i denotes some tuple of polynomials from the space. Then taking $s = e^\nu$ for sufficiently large ν ensures that the families defined by $\{P_i\}$ and $\{\tilde{P}_i\}$ have the same 0-fiber, whereas an appropriate (generic) choice of Q_i ensures that $\{\tilde{P}_i\} \in U$ for generic e .

For instance, in conjunction with the Bertini theorem, this implies that one can make the generic fiber effectively smooth.

If X is a flat family and $P \in K[\zeta]$ is such that $\{P = 0\}$ intersects X_0 properly, then the family $Y = \mathcal{F}[X \cap \{P = 0\}]$ satisfies $Y_0 = X_0 \cap \{P = 0\}$. However, if the intersection is not proper, one cannot in general predict the structure of Y_0 . The following proposition shows that under a technical modification, one can guarantee that Y_0 is given by the intersection between $\{P = 0\}$ and those components of X_0 that meet it properly.

Proposition 20. Let $X \subset N \times \mathbb{A}_K$ be a generic complete intersection and $P \in K[\zeta]$. Define Y, \tilde{Y} by

$$Y = \mathcal{F}[X \cap \{P = e^{1/\nu}\}] \quad (24)$$

$$\tilde{Y} = \mathcal{F}[(X_0 \times \mathbb{A}_K) \cap \{P = e^{1/\nu}\}] \quad (25)$$

where ν is a sufficiently large natural number.

Then $\tilde{Y}_0 = Y_0$. In particular, if $X_0 = C \cup C'$ where C' denotes the union of components of X_0 where P vanishes identically and C the rest, then $Y_0 = C \cap \{P = 0\}$.

Proof. The first part of the statement is simply Proposition 18 applied to the family $(X \times \mathbb{A}_K^s) \cap \{P = s\}$ (with e in place of t). For the second part it suffices to compute

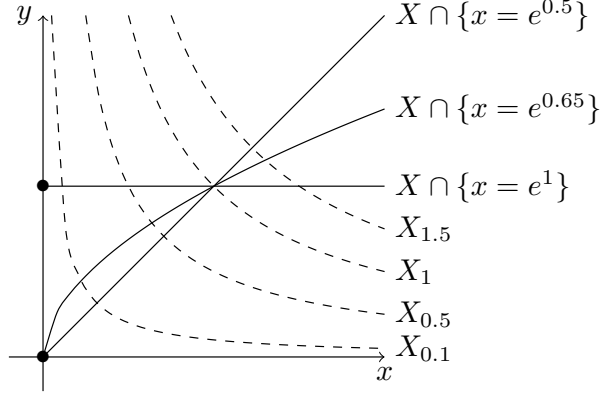


FIGURE 2. The different families in Example 21

\tilde{Y}_0 by noting that $\{P = e^{1/\nu}\}$ intersects C properly and does not intersect C' at all (in the affine space). \square

We remark that the use of the ramified factor $e^{1/\nu}$ is merely a notational convenience, to avoid reparametrizing the original family X . One could of course obtain an honest polynomial family by passing to a new deformation parameter.

Example 21. Consider $X \subset \mathbb{C}^2 \times \mathbb{C}$ given by $xy = e$, and $P = x$. Then X_0 has two components, the x and the y axes. The former intersects $\{P = 0\}$ properly at the origin, whereas the latter intersects $\{P = 0\}$ non-properly.

If we consider the family $Y = \mathcal{F}(X \cap \{P = 0\})$ we obtain $Y_0 = \emptyset$. If we consider $Y = \mathcal{F}(X \cap \{P = e\})$ we obtain $Y_0 = \{(0, 1)\}$. Finally, if we consider $Y = \mathcal{F}(X \cap \{P = e^\alpha\})$ where $0 < \alpha < 1$, then we obtain $Y = \{(0, 0)\}$, which is the intersection between the x -axis and $\{P = 0\}$ as predicted by Proposition 20.

3.2. Projections and reduction of dimension. We begin with a simple lemma on linear elimination of variables. Let $P \in K[\zeta, \zeta^{(1)}]$ be a polynomial with $\deg_{\zeta^{(1)}} P \leq 1$. The homogenization \tilde{P} of P is the polynomial in one extra variable, $\zeta_0^{(1)}, \dots, \zeta_n^{(1)}$ where the free term of P is multiplied by $\zeta_0^{(1)}$. The following lemma is immediate.

Lemma 22. Let

$$P_1, \dots, P_{n+1} \in K[\zeta, \zeta^{(1)}] \quad \Delta(P_j) \subset \Delta_j \times \Delta_{\zeta^{(1)}} \quad (26)$$

and let $\tilde{P}_1, \dots, \tilde{P}_{n+1}$ denote their homogenization, and $W \subset N \times \mathbb{P}_K^n$ the variety they define.

Let R denote the determinant of the $(n+1) \times (n+1)$ matrix whose j -th row are given by the $n+1$ coefficients of \tilde{P}_j . Then $\Delta(R) \subset \sum_{j=1}^{n+1} \Delta_j$ and $\pi(W) = \{R = 0\}$.

Lemma 23. Let

$$P_1, \dots, P_k \in K[\zeta, e] \quad (27)$$

and

$$Q_1, \dots, Q_l \in K[\zeta, \zeta^{(1)}, e] \quad \Delta(Q_j) \subset \Delta_j \times \Delta_{\zeta^{(1)}} \quad (28)$$

define a generic complete intersection $X \subset N^{(1)} \times \mathbb{A}_K$,

$$X = \mathcal{F}[\{P_1 = \dots = P_k = Q_1 = \dots = Q_l = 0\}]. \quad (29)$$

Denote by W the union of the components C of X_0 such that $\text{codim } \pi(C) > k$.

Then there exists $R \in K[\zeta, e]$ and a generic complete intersection Y ,

$$Y = \mathcal{F}[\{P_1 = \dots = P_k = R = 0\}], \quad \Delta(R) \subset \sum_{j=1}^l \Delta_j \quad (30)$$

such that $\pi(W) \subset Y_0$.

Proof. We prove the claim by reverse induction on l , starting with the case $l \geq n + 1$. In this case $W = X_0$. We replace Q_1, \dots, Q_{n+1} by their small generic perturbation (keeping the same Newton polytope). By Proposition 18 this does not change X_0 . We can thus assume without loss of generality that the generic fiber of $V \subset N \times \mathbb{P}_K^n \times \mathbb{A}_K$ cut out by P_1, \dots, P_k and the homogenizations of Q_1, \dots, Q_{n+1} has codimension $n + k + 1$. Applying Lemma 22 to Q_1, \dots, Q_{n+1} we thus obtain R giving a generic complete intersection (30). By construction the generic fiber Y_ε contains $\pi(X_\varepsilon)$, and it follows that Y_0 contains $\pi(X_0)$ as claimed.

We now consider the case $l \leq n$. Then for any component $C \subset W$, the generic fiber of $\pi : C \rightarrow \pi(C)$ has dimension at least $n + 1 - l \geq 1$. Let ℓ be a generic affine-linear form in the $\zeta^{(1)}$ variables. Then $\{\ell = 0\}$ meets X_0 properly, and taking Q_{l+1} to be ℓ we obtain a new family X' such that

$$C' := C \cap \{\ell = 0\} \subset X'_0. \quad (31)$$

Moreover,

$$\text{Clo}[\pi(C)] = \text{Clo}[\pi(C')] \quad (32)$$

since $\{\ell = 0\}$ meets the generic fiber of $\pi : C \rightarrow \pi(C)$. Since this is true for any component $C \subset W$ (with sufficiently generic ℓ), the claim follows by induction with the new family X' and $\Delta_{l+1} = \{0\}$. \square

The following lemma provides the main inductive step in our estimates.

Lemma 24. *Let $X \subset N \times \mathbb{A}_K$ given by*

$$X = \mathcal{F}[\{P_1 = \dots = P_k = 0\}], \quad P_j \in K[\zeta, e], \quad \Delta(P_j) \subset \Delta_j \quad (33)$$

be a generic complete intersection. Let $S \subset N^{(1)}$ be a variety defined by equations which are affine-linear in $\zeta, \zeta^{(1)}$. Suppose that $(X_0)^{(1)} \cap S$ does not project dominantly on any component of X_0 .

Then there exist

$$\tilde{P}_1, \dots, \tilde{P}_{k+1} \in K[\zeta, e] \quad (34)$$

with $\Delta(\tilde{P}_j) \subset \Delta_j$ for $j = 1, \dots, k$ and

$$\Delta(\tilde{P}_{k+1}) \subset (n - k + 1)\Gamma, \quad \Gamma = (n + 1)\Delta_\zeta + \sum_{i=1}^k \Delta_i \quad (35)$$

such that

$$Y = \mathcal{F}[\{\tilde{P}_1 = \dots = \tilde{P}_{k+1} = 0\}] \quad (36)$$

is a generic complete intersection and $\pi((X_0)^{(1)} \cap S) \subset Y_0 \subset X_0$.

Proof. By Proposition 18 any sufficiently small perturbation of P_1, \dots, P_k does not change X_0 . We fix such a perturbation making the generic fiber of the family effectively smooth (without changing the Newton polytopes). Without loss of generality we may assume that the family P_1, \dots, P_k is already in this form.

We let Z^0 denote the family $\tau(X)$. Recall that it is a generic complete intersection defined by the vanishing of P_1, \dots, P_k and another set of k equations Q_1, \dots, Q_k with $\Delta(Q_j) \subset (\Delta + \Delta_\zeta) \times \Delta_{\zeta(1)}$. By Corollary 17 every component of $(X_0)^{(1)}$ is also a component of Z_0^0 . Moreover, any extra components do not project dominantly on a component of X_0 by Proposition 15.

For $j = 1, \dots, n - k + 1$ we define Z^j to be the family obtained from Z^{j-1} by intersecting with a generic linear combination \tilde{Q}^j of the given equations of S , using Proposition 20. We define R^j, Y^j by applying Lemma 23 to Z^j . We have $\Delta(R_j) \subset \Gamma$. Finally, we take $\tilde{P}_{k+1} = R_1 \cdots R_{n-k+1}$ which gives $Y = Y^1 \cup \dots \cup Y^{n-k+1}$.

Let C be a component of $(X_0)^{(1)}$. Then it is also a component of Z_0^0 . Since C projects dominantly on a component of X_0 , it cannot be contained in S . Hence by genericity $C' := C \cap \{\tilde{Q}^1 = 0\}$ is strictly contained in C , and by Proposition 20 we have $C' \subset Z_0^1$. Moreover, clearly $C' \cap S = C \cap S$. Thus it will suffice to prove that

$$\pi(Z_0^1 \cap S) \subset Y_0. \quad (37)$$

Now let C be any component of Z_0^1 . If $\text{codim}(\pi(C)) > k$ then

$$\pi(C \cap S) \subset \pi(C) \subset Y_0^1 \subset Y_0 \quad (38)$$

where the middle inclusion follows from Lemma 23. Otherwise, C projects dominantly on a component of X_0 , so it again cannot be contained in S . Again by genericity $C' := C \cap \{\tilde{Q}^2 = 0\}$ is strictly contained in C , and by Proposition 20 we have $C' \subset Z_0^2$. Moreover, clearly $C \cap S = C' \cap S$. Thus it will suffice to prove that

$$\pi(Z_0^2 \cap S) \subset Y_0. \quad (39)$$

We repeat this process until Z_0^{n-k+1} . At this point every component has dimension at most $n - k - 1$ even before the π projection, and thus project to a set of codimension greater than k . As in (38) it then follows that

$$\pi(C \cap S) \subset \pi(C) \subset Y_0^{n-k+1} \subset Y_0 \quad (40)$$

for every component C of Z_0^{n-k+1} , thus concluding the proof. \square

Remark 25. If the polytopes $\Delta_1, \dots, \Delta_k$ are co-ideal polytopes (see Remark 4) then they are stable under differentiations. In this case in the definition of Γ in (35) one can replace $(n+1)\Delta_\zeta$ by $(n-k+1)\Delta_\zeta$. For such polytopes, all of our main results remain valid with this refined definition of Γ .

4. PROOFS OF THE MAIN ESTIMATES

In this section we prove our main results on the number of isolated solutions of a differential system of equations and the degree of its reduction. We begin with a lemma demonstrating the application of Lemma 24 to this context.

Lemma 26. Let $X \subset \Omega^{(l)} \times \mathbb{A}_K$ be a generic complete intersection family, and suppose that $(X_0)^{(1)} \cap \Xi$ does not project dominantly on any component of X_0 . Then the family Y defined by applying Lemma 24 to X with $M = \Omega^{(l)}$ and $S = \Xi$ satisfies $X_0 \cap \mathcal{J}^l(M) = Y_0 \cap \mathcal{J}^l(M)$.

Proof. Since $Y_0 \subset X_0$, one inclusion is obvious. In the other direction, let $x^{(l)} \in X_0$. Then $x^{(l)(1)} \in (X_0)^{(1)} \cap \Xi$ so by Lemma 24 $x^{(l)} = \pi(x^{(l)(1)}) \in Y_0$ as well. \square

We now prove our result for the case that Y is a fiber X_0 of a generic complete intersection family X . In particular the following proposition implies (and generalizes) Theorem 5.

Proposition 27. *Let $X \subset \Omega^{(l)} \times \mathbb{A}_K$ be a generic complete intersection family defined by polynomials with Newton polytopes $\Delta_1, \dots, \Delta_k$, and suppose that X_0 admits finitely many solutions. Denote*

$$\Gamma = s\Delta_{\xi}^{(l)} + \Delta_1 + \dots + \Delta_k. \quad (41)$$

Then

$$\mathcal{N}(X_0) \leq C_{s,k} V(\Delta_1, \dots, \Delta_k, \Gamma, \dots, \Gamma), \quad C_{s,k} = s!(\delta + 2)^{\delta(\delta+1)/2} \quad (42)$$

where $\delta := s - k$.

Proof. Let $X(0) = X$. As long as $X(j)_0$ has positive dimension, we define $X(j+1)$ to be the family obtained by applying Lemma 24 to $X(j)$ with the ambient space $\Omega^{(l)}$ and $S = \Xi$. The lemma is applicable by application of Lemma 12, since $X(j)_0$ admits finitely many solutions. We denote the extra equation obtained in this process¹ by R_{j+1} . By Lemma 26 we have $\mathcal{N}(X(j+1)_0) = \mathcal{N}(X(j)_0)$.

The Newton polytope of R_j is contained in $(\delta + 2)^j \Gamma$, as one can see by the following simple induction:

$$\Delta(R_j) \subset (\delta + 1) \left(\sum_{i=0}^{j-1} (\delta + 2)^i \Gamma \right) \subset (\delta + 1) \frac{(\delta + 2)^j - 1}{\delta + 1} \Gamma \subset (\delta + 2)^j \Gamma \quad (43)$$

Eventually we obtain the zero-dimensional variety $X(\delta)$. From the above we conclude that $\mathcal{N}(X_0) = \mathcal{N}(X(\delta)_0)$, which is certainly bounded by the number of points in $X(\delta)_0$. Since the number of points of a flat limit is certainly bounded by the number of points of the generic fiber, we have by the BKK theorem

$$\mathcal{N}(X_0) \leq s! V(\Delta_1, \dots, \Delta_k, (\delta + 2)^1 \Gamma, \dots, (\delta + 2)^{\delta} \Gamma) \quad (44)$$

$$= (\delta + 2)^{\delta(\delta+1)/2} s! V(\Delta_1, \dots, \Delta_k, \Gamma, \dots, \Gamma) \quad (45)$$

as claimed. \square

We now prove our result for arbitrary varieties $Y \subset \Omega^{(l)}$.

Proof of Theorem 6. Let $X(k) \subset \Omega^{(l)} \times \mathbb{A}_K$ be the (constant) family cut out by the given equations with Newton polytopes $\Delta_1, \dots, \Delta_k$.

For $j \geq k$, let P_j denote a generic linear combination of the given equations with Newton polytope Δ defining Y . Define $X(j+1)$ to be the family obtained by application of Proposition 20 to $X(j)$ and P_j . Write $X(j)_0 = A(j) \cup B(j)$, where $A(j)$ denotes the union of the components of $X(j)_0$ that are contained in Y , and $B(j)$ the rest.

The number of solutions of Y which are contained in $C(j) = A(j) \setminus B(j)$ is bounded by

$$\mathcal{N}(C(j)) \leq C_{s,k} V(\Delta_1, \dots, \Delta_j, \Gamma_j, \dots, \Gamma_j). \quad (46)$$

Indeed, $C(j)$ is the flat limit of the family $X(j)$ in the ambient space $\Omega^{(l)} \setminus B(j)$, and the bound thus follows from Proposition 27. The claim of the theorem will follow once we show that $Y = C(k) \cup \dots \cup C(s)$.

¹In fact each application of Lemma 24 perturbs all the equations defining $X(j)$; but the Newton polytopes remain unchanged.

Let $x \in Y$, and we will show that it belongs to some $C(j)$. Certainly $x \in X(k)_0$. If $X \not\subset B(k)$, we are done. Otherwise x is contained in some component $G \subset X(k)_0$ with $G \not\subset Y$, and we may assume by genericity that P_k does not vanish identically on G . Then according to Proposition 20, $X(k+1)_0$ contains $G \cap \{P_k = 0\}$, and in particular $x \in X(k+1)_0$.

We continue in the same manner. The process must stop at $j = s$ (if not before), because at this point $X(j)_0$ consists of isolated points, so $x \in X(j)_0$ implies $x \in C(j)$. \square

Finally we prove the more general Corollary 8.

Proof of Corollary 8. We indicate the minor changes required in the proofs of Proposition 27 and Theorem 6.

The proof of Proposition 27 carries out verbatim as long as $X(j)_0$ does not have a component which is a component of $\mathcal{R}(Y)$. Let $S(j) \subset \mathcal{R}(Y)$ denote the components of codimension j . Then these are also components of $X(j)_0$, and their degrees can be bounded by the BKK theorem. We then define $X(j+1)$ in the same way, but restricting the ambient space to $\Omega^{(l)} \setminus (S(1) \cup \dots \cup S(j))$. This insures that Lemma 24 is again applicable. It remain only to note that the degree of the remaining components of $\mathcal{R}(Y)$ in the new ambient space is the same as the degree in the original ambient space. The proof is thus concluded by induction. The resulting bound has an extra multiplicative factor of $s - k + 1$ corresponding to the fact that we separately bound the degree in each dimension².

We can now carry out the proof of Theorem 6 in the same way, noting that since $Y = C(k) \cup \dots \cup C(s)$, any component of $\mathcal{R}(Y)$ must be a component of $\mathcal{R}(C(j))$ for some $j = k, \dots, s$. \square

5. DIOPHANTINE APPLICATIONS

In the papers [8, 6] the effective estimate of Theorem 2 has been used to derive estimates for some counting problems of a diophantine nature. In this section we illustrate our result by improving the estimates presented in these papers.

In this section we assume that $K = \mathbb{C}$, and the differentiation operator D is chosen such that the field of constants is $k = \bar{\mathbb{Q}}$.

5.1. Transcendental points in subvarieties of semi-abelian varieties. Recall that a semi-abelian variety is an extension of an abelian variety by a torus. Let A be a semi-abelian variety and $\Gamma \subset A$ as subgroup of finite rational rank $r := \dim_{\mathbb{Q}} \Gamma \otimes \mathbb{Q}$. Finally let X be a subvariety of A .

In [8] effective bounds are given on the number of points in the intersection $X \cap \Gamma$ under various conditions on Γ, X . The estimates are presented in terms of the following data. Suppose that A has dimension n and is defined over k . We assume that A is embedded as a locally closed subset of a projective space $\mathbb{C}P^N$. Let $\omega_1, \dots, \omega_n$ be a basis of translation-invariant differential forms on A .

We assume that A is covered by t affine charts, and that each ω_i is given in each chart by a polynomial in the local coordinates x_1, \dots, x_N and dx_1, \dots, dx_n . Let d_A be the maximal degree of the equations defining A in any of the charts, and d_ω be the maximal degree of any of the polynomials defining $\omega_1, \dots, \omega_n$ in any of the

²one could obviously obtain a sharper estimate taking into account the different BKK estimate for each dimension.

charts. Finally let d_X denote the maximal degree of the equations defining X in any of the charts. We assume for simplicity of the formulation that $d_X \geq d_A$.

The following is a refined form of the main result of [8].

Theorem 28 (cf. [8, Theorem 1.1]). *Suppose A, X are defined over k and there exist no positive-dimensional subvarieties $X_1, X_2 \subset A$ such that $X_1 + X_2 \subset X$. Then*

$$\#[(X \cap \Gamma) \setminus X(k)] \leq F_{N,n,r} \cdot t \cdot d_\omega^{Nr} d_x^n, \quad (47)$$

where

$$F_{N,n,r} = \frac{E_{N(r+1), N-n}}{N(r+1)!} \binom{Nr+n}{n} d_A^{N-n} 2^n. \quad (48)$$

Proof. In [8] it is shown that Γ is contained in a finite-dimensional differential algebraic subgroup G of A such that $B := (X \cap G) \setminus X(k)$ is finite. It is therefore enough to bound the number of points in B . Moreover, it is shown that in each of the affine charts on A , the group G can be written in the form

$$G = \{x : x^{(r)} \in S\}$$

where $S \subset (K^N)^{(r)}$ is a variety defined by equations of degrees bounded by d_ω .

We will apply Theorem 6 with parameters

$$\begin{aligned} M &= K^N, \\ l &= r, \\ \Omega^{(r)} &= M^{(r)} \setminus \{\xi^{(1)} = 0\}, \\ Y &= \Omega^{(r)} \cap S \cap \pi^{-1}(X). \end{aligned}$$

Since Y is contained in $\pi^{-1}(A)$ which has codimension $N - n$ and is cut out by equations of degree at most d_A , we can apply Theorem 6 with

$$\Delta_1 = \cdots = \Delta_{N-n} = d_A \Delta_\xi, \quad \Delta = d_X \Delta_\xi + d_\omega \Delta_\xi^{(r)}. \quad (49)$$

We note that $\Delta \subset \Delta' + \Delta''$ where $\Delta' = 2d_X \Delta_\xi$ and Δ'' is the polytope of polynomials of degree d_ω in $\xi^{(1)}, \dots, \xi^{(r)}$.

Finally, Y admits finitely many solutions, and by Theorem 6 we have

$$\begin{aligned} N(Y) &\leq E_{s, N-n} V(\Delta_1, \dots, \Delta_{N-n}, \Delta' + \Delta'', \dots, \Delta' + \Delta'') \\ &= E_{s, N-n} \binom{s - N + n}{n} V(\underbrace{d_A \Delta_\xi}_{N-n \text{ times}}, \underbrace{\Delta'}_{n \text{ times}}, \underbrace{\Delta''}_{s-N \text{ times}}) \\ &\leq E_{s, N-n} \binom{s - N + n}{n} d_A^{N-n} (2d_X)^n d_\omega^{N-n} V(\Delta_\xi^{(l)}, \dots, \Delta_\xi^{(l)}) \\ &= E_{s, N-n} \binom{s - N + n}{n} 2^n (s!)^{-1} d_A^{N-n} d_\omega^{s-N} d_X^n \end{aligned}$$

where in the second equality we expand the mixed volume by multi-linearity and use the fact that the value is non-zero if and only if Δ' appears exactly n times. \square

We note in particular that the bound of Theorem 28 is singly exponential with respect to N and r , and has the natural asymptotic d_X^n with respect to d_X . We also remark that since A is smooth, the arguments of this paper could have been applied with small changes in the ambient space A rather than K^N , leading to somewhat better estimates.

We next present a version of Theorem 28 for a torus. This result may be seen as an analog of the Kushnirenko theorem. Similar analogs of the BKK theorem involving mixed volumes can be obtained in a similar manner.

Theorem 29. *Let $A = (\mathbb{C}^*)^n$ and let $X \subset A$ be defined over k by equations with Newton polytope Δ . Suppose that there exist no positive-dimensional subvarieties $X_1, X_2 \subset A$ such that $X_1 + X_2 \subset X$. Then*

$$\#[(X \cap \Gamma) \setminus X(k)] \leq F_{2n,n,r} 2^{n(2r+1)} \text{Vol}(\Delta) \quad (50)$$

Proof. We embed A in $\mathbb{C}^{2n} = \mathbb{C}[x_1, \dots, x_n, y_1, \dots, y_n]$ with the equations $x_i y_i = 1$. Then the invariant forms dx_i/x_i become polynomials of degree 2, so $d_A = d_\omega = 2$. The proof proceeds in a manner analogous to the proof of Theorem 28. \square

A special case with $n = 2$ in Theorem 29 is of some particular interest. It has been considered by Buium [5] who gave an iterated-exponential bound, answering a question posed by Bombieri. The bound was improved in [8] to a doubly-exponential bound as follows.

Theorem 30 ([8, Corollary 1.2]). *Let $f \in \bar{\mathbb{Q}}[x, y]$ be an irreducible polynomial of degree d , whose zero-locus in $(\mathbb{C}^*)^2$ is not a translate of a torus. Let $\Gamma \subset \mathbb{C}^*$ be a subgroup of rational rank r . Then*

$$\#\{(x, y) \in \Gamma^2 : f(x, y) = 0, x \notin \bar{\mathbb{Q}}\} \leq d^{r^2} (r+1)^{2(2^r+1)} \quad (51)$$

From Theorem 29 we obtain

Theorem 31. *Let $f \in \bar{\mathbb{Q}}[x, y]$ be an irreducible polynomial with Newton polygon Δ , whose zero-locus in $(\mathbb{C}^*)^2$ is not a translate of a torus. Let $\Gamma \subset \mathbb{C}^*$ be a subgroup of rational rank r . Then*

$$\#\{(x, y) \in \Gamma^2 : f(x, y) = 0, x \notin \bar{\mathbb{Q}}\} \leq F_{4,2,r} 2^{4r+2} \text{Vol}(\Delta). \quad (52)$$

We note in particular that the bound of Theorem 30 is singly-exponential in r , and if we assume that $\deg f = d$ then $\text{Vol}(\Delta) = d^2/2$.

We also present a refined form of [8, Theorem 1.1], replacing the condition on X in Theorem 29 by a condition on the lattice Γ . For simplicity we present the result in the context of the torus, although the proof works in general in the same way as the proof of 28.

Theorem 32. *Let $A = (\mathbb{C}^*)^n$ and let $X \subset A$ be defined by equations with Newton polytope Δ , and suppose that it does not contain a translate of a non-trivial subtorus. Suppose that $\Gamma \cap A(k) = \{0\}$. Then*

$$\#(X \cap \Gamma) \leq n(2r+1) F_{2n,n,r} 2^{n(2r+1)} \text{Vol}(\Delta) \quad (53)$$

Proof. In [8, Lemma 6.1] it is shown that Γ is contained in a finite-dimensional differential algebraic subgroup G of $(K^*)^n$ such that $X \cap G$ intersects only finitely many translates $c_1(k^*)^n, \dots, c_\nu(k^*)^n$ of $(k^*)^n$. By the assumption on Γ , each such translate intersects Γ at most once, so it will suffice to give a bound on ν .

We choose $M, \Omega^{(r)}$ as in the proof of Theorem 28 and write $X \cap G$ as the set of solutions of a variety $Y \subset \Omega^{(r)}$. The logarithmic derivative $l(x) = (Dx)/x$ takes a different value $l(c_j)$ on each of the translates $c_j(k^*)^n$, so the solutions of Y are contained in a union of ν disjoint planes $\xi^{(1)} = c_j \xi$ in $\Omega^{(r)}$ (and meets each of them). Then the same is true for the Zariski closure $\mathcal{R}(Y)$, and it is thus enough to bound the degree of this variety. The result now follows by Corollary 8. \square

5.2. Isogeny classes of elliptic curves. Denote by E_x the elliptic curve with j -invariant $x \in \mathbb{C}$. We write $E_x \sim E_y$ if E_x is isogenous to E_y . We denote

$$\text{Iso}(x) = \{y \in \mathbb{C} : E_y \sim E_x\} \quad (54)$$

and more generally, for $\bar{x} \in \mathbb{C}^n$,

$$\text{Iso}(\bar{x}) := \{\bar{y} \in \mathbb{C}^n : E_{x_j} \sim E_{y_j} \text{ for } j = 1, \dots, n\}. \quad (55)$$

In [6] the authors, following a question of Mazur, consider the following problem: given an automorphism α of \mathbb{CP}^1 and $\tau \in \mathbb{C}$, estimate the number of elliptic curves E_ρ satisfying $E_\rho \sim E_\tau$ and $E_{\alpha \cdot \rho} \sim E_{\alpha \cdot \tau}$. In particular they obtain the following estimate.

Theorem 33 ([6, Section 6.1]). *Assume τ is transcendental and set*

$$S := \{(z, w) : w = \alpha \cdot z\} \cap \text{Iso}(\tau, \alpha \cdot \tau). \quad (56)$$

Then the size of S , i.e. the number of elliptic curves E_ρ satisfying $E_\rho \sim E_\tau$ and $E_{\alpha \cdot \rho} \sim E_{\alpha \cdot \tau}$, is at most³ $2^{24}36^7 \simeq 10^{18}$.

The proof of this theorem, as in §5.1, is based on the introduction of a differential algebraic construction. Namely, recall that the Schwratzian operator is defined by

$$S(x) = \left(\frac{x''}{x'}\right)' - \frac{1}{2} \left(\frac{x''}{x'}\right)^2 \quad (57)$$

where we interpret this as an operator on our differentially closed field \mathbb{C} , and write x' as a shorthand for Dx . We introduce the differential operator

$$\chi(x) = S(x) + R(x)(x')^2, \quad R(x) = \frac{x^2 - 1968x + 2654208}{2x^2(x - 1728)^2} \quad (58)$$

which is a third order algebraic differential operator vanishing on Klein's j -invariant j [4].

Let $\tau \in \mathbb{C}$ be transcendental (i.e., non-constant with respect to our chosen differential structure). By a theorem of Buim [4], the set $\chi^{-1}(\chi(\tau))$ is the Kolchin closure of $\text{Iso}(\tau)$. One may therefore attempt to study the set S in (56) by considering the (possibly larger) set

$$Z := \{(z, w) : w = \alpha \cdot z\} \cap \{(z, w) : \chi(z) = \chi(\tau), \chi(w) = \chi(\alpha \cdot \tau)\}. \quad (59)$$

Of course, even if the set S is finite (which is not an obvious assertion), it is not clear a-priori that the set Z must also be finite. However, in [6, Section 6.1] it is proven that this is indeed the case. Since Z is a set given by differential algebraic conditions, it is thus possible to estimate its size using the methods of [8], and this is carried out in [6] to give Theorem 33.

We now apply our results, specifically Theorem 5, to estimate the size of Z . We begin by expressing Z as the set of solutions of a variety of differential conditions Y . Let $\alpha \cdot z = \frac{az+b}{cz+d}$. We choose $M = \mathbb{C}^2, l = 3$ and let $\Omega^{(l)}$ be the open dense subset of $M^{(l)}$ obtained by removing the polar divisor of $\alpha \cdot \xi$ (i.e. $\{c\xi + d = 0\}$) as well as the polar divisors of $\chi(\xi)$ and $\chi(\eta)$.

We write six explicit equations for Y . The first is given by $P_1 : (c\xi + d)\eta = a\xi + b$. We obtain the next three equations $P_{2, \dots, 4}$ by taking the first three derivatives of P_1 with respect to D (and replacing $D\xi$ by $\xi^{(1)}$, etc.). One easily checks that these

³The constant appearing in [6] contains a minor computational error, this is a tentative correction.

equations define a complete intersection (in fact, they define the third prolongation of the graph of α : at this point it is essential that we removed the polar divisor $c\xi + d = 0$). Clearly $\Delta(P_j) \subset \Delta_\xi^{(3)} + \Delta_\eta^{(3)}$ for $j = 1, 2, 3, 4$.

The next two equations P_5, P_6 are given by $\chi(\xi) = \chi(\tau)$ and $\chi(\eta) = \chi(\alpha \cdot \tau)$ respectively, where we clear out all the denominators. We have $\Delta(P_5) \subset 6\Delta_\xi^{(3)}$ and $\Delta(P_6) \subset 6\Delta_\eta^{(3)}$. Since these equations are linear in $\xi^{(3)}$ and $\eta^{(3)}$ respectively, it is not hard to check that they are each irreducible (this was already noted in [6, Section 5.2]). It follows that $P_{1\dots 6}$ define a complete intersection. Indeed, otherwise P_1, \dots, P_5 would imply P_6 . But P_5 , being an equation of order 3, admits infinitely many solutions z , and for each of these $(z, \alpha \cdot z)$ would be a solution of $P_{1,\dots,5}$ and hence also of P_6 , contradicting the fact that Y has finitely many solutions.

Finally, we apply Theorem 5 to Y . In computing Γ we also take into account Remark 25,

$$\Gamma \subset [(8 - 6 + 1) + 4 + 6](\Delta_\xi^{(3)} + \Delta_\eta^{(3)}) = 13(\Delta_\xi^{(3)} + \Delta_\eta^{(3)}) \quad (60)$$

and

$$\begin{aligned} \mathcal{N}(Y) &\leq 4^3 \cdot 8! V(\underbrace{\Delta_\xi^{(3)} + \Delta_\eta^{(3)}}_{4 \text{ times}}, 6\Delta_\xi^{(3)}, 6\Delta_\eta^{(3)}, \underbrace{13(\Delta_\xi^{(3)} + \Delta_\eta^{(3)})}_{2 \text{ times}}) \\ &\leq 2^6 6^2 13^2 \cdot 8! V(\Delta_\xi^{(3)}, \Delta_\eta^{(3)}, \underbrace{\Delta_\xi^{(3)} + \Delta_\eta^{(3)}}_{6 \text{ times}}) \\ &= 2^6 6^2 13^2 \binom{6}{3} \cdot 8! V(\underbrace{\Delta_\xi^{(3)}}_{4 \text{ times}}, \underbrace{\Delta_\eta^{(3)}}_{4 \text{ times}}) \\ &= 2^6 6^2 13^2 \binom{6}{3} = 2^{10} \cdot 3^3 \cdot 13^2. \end{aligned}$$

In conclusion,

Theorem 34 (cf. Theorem 33). *Assume τ is transcendental. Then the number of elliptic curves E_ρ satisfying $E_\rho \sim E_\tau$ and $E_{\alpha \cdot \rho} \sim E_{\alpha \cdot \tau}$, is at most $2^{10} \cdot 3^3 \cdot 13^2 \simeq 5 \times 10^6$.*

Remark 35. *One could have derived a bound directly using Theorem 6 without the derivation of the extra equations $P_{2,\dots,4}$ and the fact that the intersection with P_5 and P_6 is complete. We presented this more detailed approach because it gives a somewhat better estimate, and also to illustrate a computation involving mixed volumes in Theorem 5.*

The computation above could certainly be improved somewhat: by using the precise Newton polytopes of $P_{1,\dots,6}$; by accurately computing the resulting mixed volumes; and by following the proof of Theorem 5 where various inaccurate estimates were invariably made.

Generalizing to varieties of arbitrary dimension and degree, [6] gives the following result.

Corollary 36 ([6, Corollary 6.9]). *Let $V \subset \mathbb{C}^n$ be a Zariski closed set of dimension m and $\bar{\tau} \in \mathbb{C}^n$ an n -tuple of transcendental numbers. Let W denote the Zariski closure of $V \cap \text{Iso}(\bar{\tau})$. Then W is a weakly-special variety, and³*

$$\deg W \leq (2^n \deg V)^{3 \cdot 2^{3m}} 6^{2^{3m}-1}. \quad (61)$$

We refer the reader to [6] for the definition of a weakly special variety. As for the degree estimate, the proof of this corollary proceeds in a manner analogous to the proof of Theorem 33. Arguing in a manner analogous to the proof of Theorem 34 and using Corollary 8 we obtain the following result.

Corollary 37. *Let $V \subset \mathbb{C}^n$ be a Zariski closed set defined by equations of degree d , and $\bar{\tau} \in \mathbb{C}^n$ an n -tuple of transcendental numbers. Let W denote the Zariski closure of $V \cap \text{Iso}(\bar{\tau})$. Then $\deg W \leq G_n d^n$ where G_n is an explicit constant, singly-exponential in n .*

In conclusion we remark the estimates in terms of degrees in this section could be generalized to estimates in terms of volumes of Newton polytopes, with the proofs extending verbatim.

REFERENCES

- [1] D. N. Bernstein. The number of roots of a system of equations. *Funkcional. Anal. i Priložen.*, 9(3):1–4, 1975.
- [2] A. Buium. Intersections in jet spaces and a conjecture of S. Lang. *Ann. of Math. (2)*, 136(3):557–567, 1992.
- [3] Alexandru Buium. Effective bound for the geometric Lang conjecture. *Duke Math. J.*, 71(2):475–499, 1993.
- [4] Alexandru Buium. Geometry of differential polynomial functions. III. Moduli spaces. *Amer. J. Math.*, 117(1):1–73, 1995.
- [5] Alexandru Buium. Uniform bound for generic points of curves in tori. *J. Reine Angew. Math.*, 469:211–219, 1995.
- [6] James Freitag and Tom Scanlon. Strong minimality and the j -function. *Preprint*, arXiv:1402.4588, 2014.
- [7] Heisuke Hironaka. Stratification and flatness. In *Real and complex singularities (Proc. Ninth Nordic Summer School/NAVF Sympos. Math., Oslo, 1976)*, pages 199–265. Sijthoff and Noordhoff, Alphen aan den Rijn, 1977.
- [8] Ehud Hrushovski and Anand Pillay. Effective bounds for the number of transcendental points on subvarieties of semi-abelian varieties. *Amer. J. Math.*, 122(3):439–450, 2000.
- [9] E. R. Kolchin. *Differential algebra and algebraic groups*. Academic Press, New York-London, 1973. Pure and Applied Mathematics, Vol. 54.
- [10] A. G. Kušnirenko. Newton polyhedra and Bezout’s theorem. *Funkcional. Anal. i Priložen.*, 10(3, 82–83.), 1976.
- [11] Anand Pillay. Model theory and Diophantine geometry. *Bull. Amer. Math. Soc. (N.S.)*, 34(4):405–422, 1997.
- [12] Thomas Scanlon. Diophantine geometry from model theory. *Bull. Symbolic Logic*, 7(1):37–57, 2001.
- [13] A. Seidenberg. Abstract differential algebra and the analytic case. *Proc. Amer. Math. Soc.*, 9:159–164, 1958.
- [14] A. Seidenberg. Abstract differential algebra and the analytic case. II. *Proc. Amer. Math. Soc.*, 23:689–691, 1969.

UNIVERSITY OF TORONTO, TORONTO, CANADA
E-mail address: galbin@gmail.com